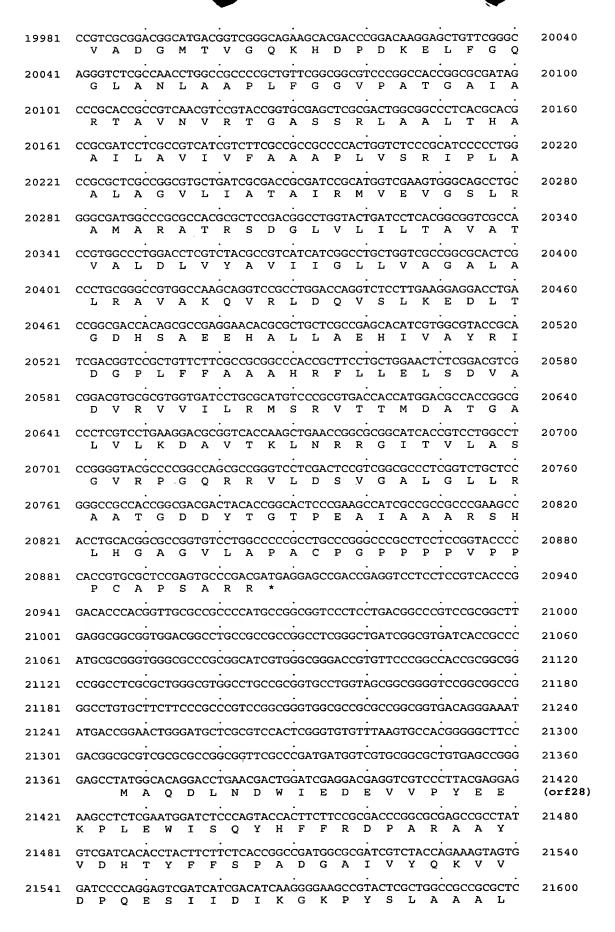
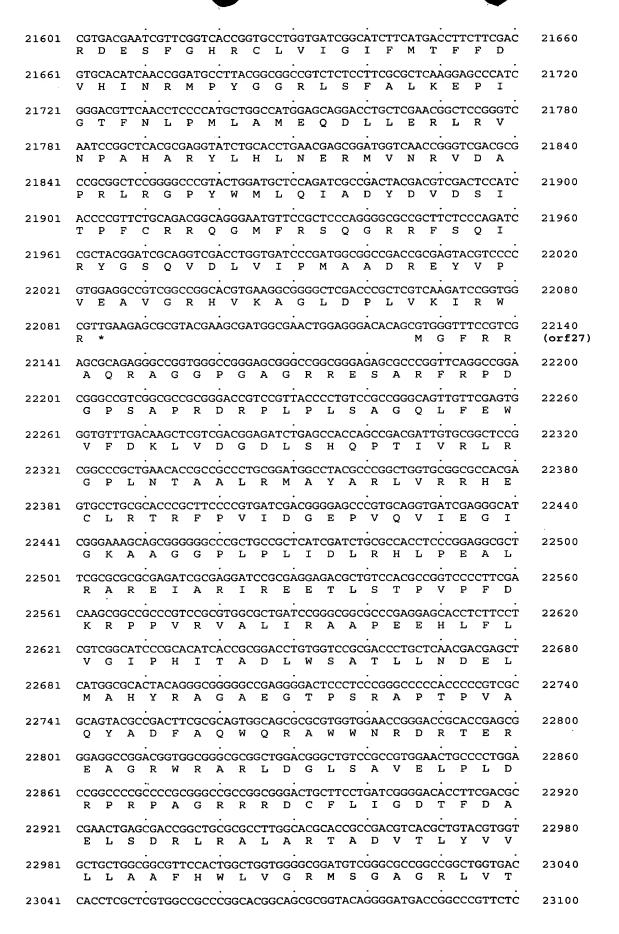
## SEQUENCE LISTING

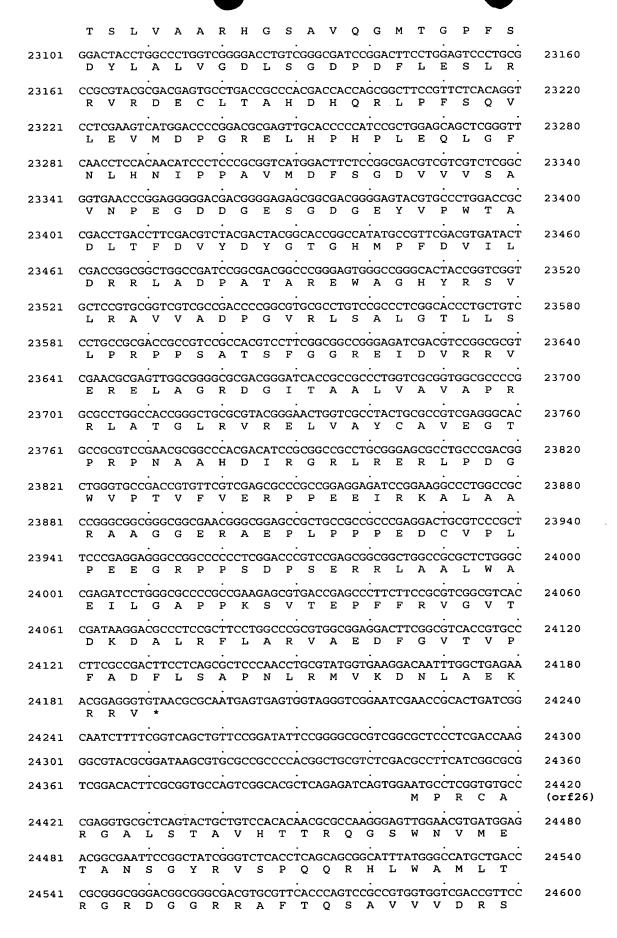
## SEQ ID NO: 1 BLM gene cluster ORFS 30 through 8

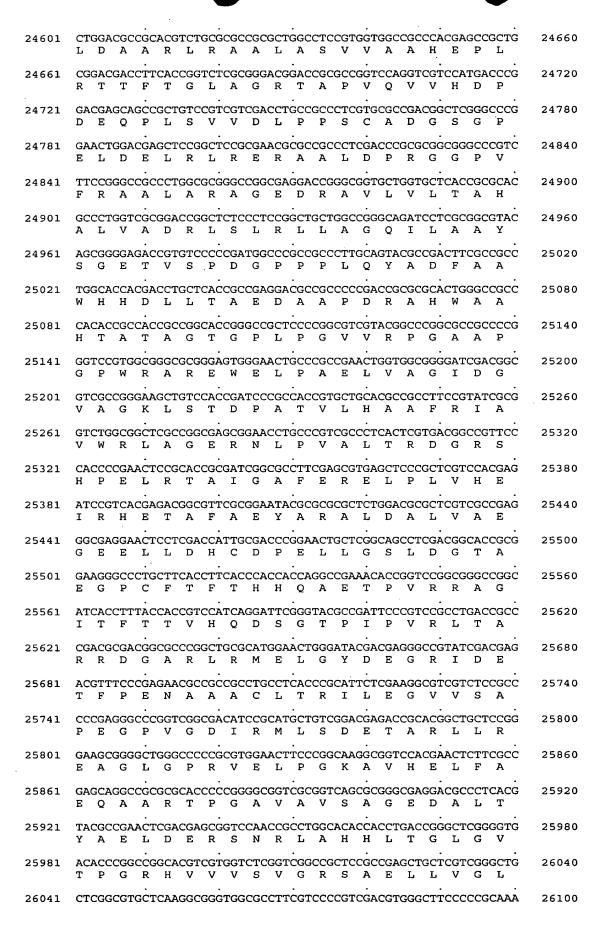
(note orf 31-40 on sequence 1-18660 are translated on the reverse strand and on a separate file)

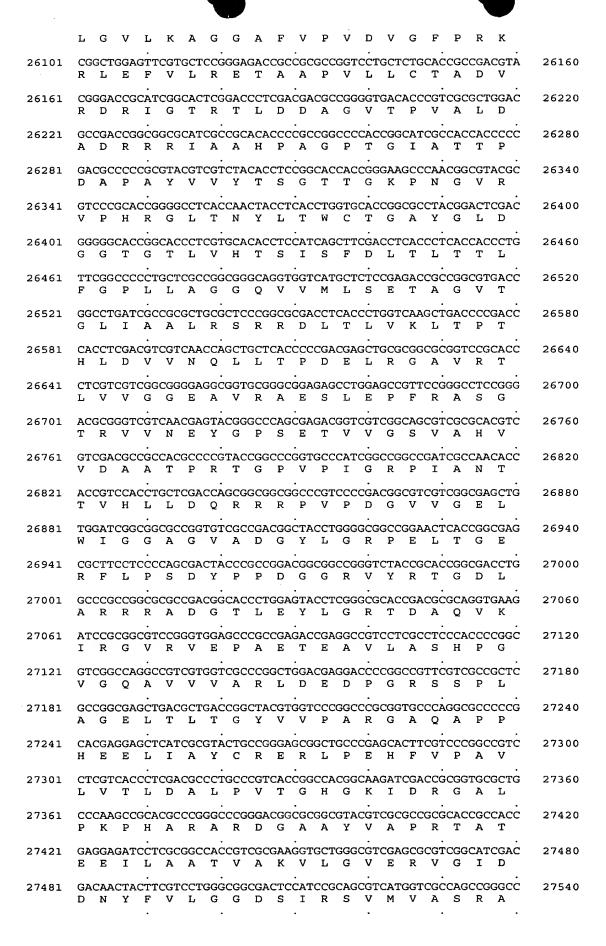
separate	Tile)	
18601	ACCCATCTCATAGGTGTACGCGCTGGAGCATTCGGGGCACGACGGAAGGTTCTCGGTCAC	18660
18661	GAGAGCACTGTAAGCCCGAACCCGCAAGCATGACGAATTGCAAAATTGTGCAAGTCGCTA	18720
18721	CATGATGGTCCGGCTGTGCCCGCAGGTAGCCGCGGGCACAGCACCAGACGCTGCCTCCGC	18780
18781	GCACCGCGGGGAGGCCCGGTGAGGCGAGGCTGAGGTTCCGCTGTAT  M P V P L Y	18840 (orf30)
18841	CAGGCGAAGGCCGAGTTCTTCCGGATGCTGGGGCACCCGGTCCGCATCCGCGTACTGGAG Q A K A E F F R M L G H P V R I R V L E	18900
18901	CTGCTGCAGGACGGCCGATGCCGGTGCGTGATCTGCTGCGGCGATCGAGATCGAGCCC L L Q D G P M P V R D L L A A I E I E P	18960
18961	TCGGCGCTGTCCCAGCAGCTGGCGGTGTTGCCCCGCTCGGGCATCGTGACCTCCACCCGC S A L S Q Q L A V L R R S G I V T S T R	19020
19021	ACGGGTTCCACGGTCGTCTACGAGCTGGCCGTGGCGAGCTGATGTCCGCC T G S T V V Y E L A G G D V A E L M S A	19080
19081	GCGCGCCGCATCCTGACCGAGATGCTCAATGGCCAGCACGAGCTGCTGAGGAGCTGAGGAGAGCTGAGAGAAGAAGAAAAAAAA	19140
19141	GAAGCCGAGGTCAGTGCCCGGTGAGCTCCCTCGCCGTCCGGGTGGGAGCCCGGGTGCGTT E A E V S A R *	19200
	M S S L A V R V .G A R V R S	(orf29)
19201	CCGTGCTGCCCACCCGCGCGACCTCGCGGGCATGGGCCGCGACGTGATCTAC V L P T R A D L A G M G R S P R R D L L	19260
19261	TGGCCGGTCTGACCGTGGCGATCGTGGCCCTCGCGCCTCGGATTCGGCGTCTCCT A G L T V A I V A L P L A L G F G V S S	19320
19321	CCGGTCTCGGCGCGGGGCGCGCGGCGCGCGCGGGGGGCGCGCGGGGG	19380
19381	TATTCGGTGGGTCGAATCTCCAGGTGTCCGGGCCCACGGGCGCCATGACCGTGGTCCTGG F G G S N L Q V S G P T G A M T V V L V	19440
19441	TGCCCATCGTCGCCGGTACGGCCCCGGCGGTGTCCTCACGGTCGGCCTGCTCGCCGGAC P I V A R Y G P G G V L T V G L L A G L	19500
1/9501	TGATGCTGATCGCGCTCGCCCGCCGCCGGCCGCTACATGCAGTACGTGCCGGCCC	19560
19561	CGGTGGTGGAGGGCTTCACCCTCGGCATCGCCTGCGTGATCGGCTTGCAGCAGGTGCCGA V V E G F T L G I A C V I G L Q Q V P N	19620
19621	ACGCCTGGGAGTCGCCAAGCCGGAGGGCGACAAGGTCCTCGTCGTGACCTGGCGCGCGGAA L G V A K P E G D K V L V V T W R A V	19680
19681	TCGAGACCTTCGCCGGGGCGCCCAACTGGACCGCTGCCGGACTGGCGGCAGCGGTCGCCG	19740
19741	CGGTCATGCTGACCGGCGCGGCGGTGGCGGCCGGTCGTTCTCCCTCC	19800
19801	CCGGTGCCACCGTCGTGGCCACCTGGACGCGGCCCGATCGGGGACC G A T V V A Q L C H L D A A R P I G D L	19860
19861	TGCCCGCGGGGCTGCCCGCCCGTCGCTTCCTGGACCTCGGAGCACTGGGCTCGC PAGLPAPSLAFLDLGALGSL	19920
19921	TGCTGGCGCCTGCCGTGGCGGCCCTTGCCGCGTTGGAATCGCTGCTGTCGGCGT L A P A V A V A A L A A L E S L L S A S	19980

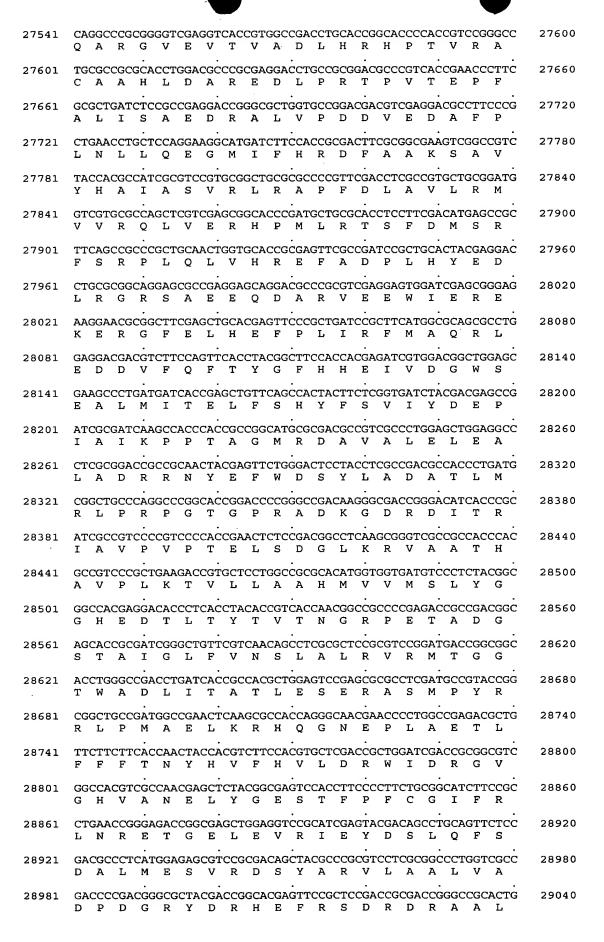


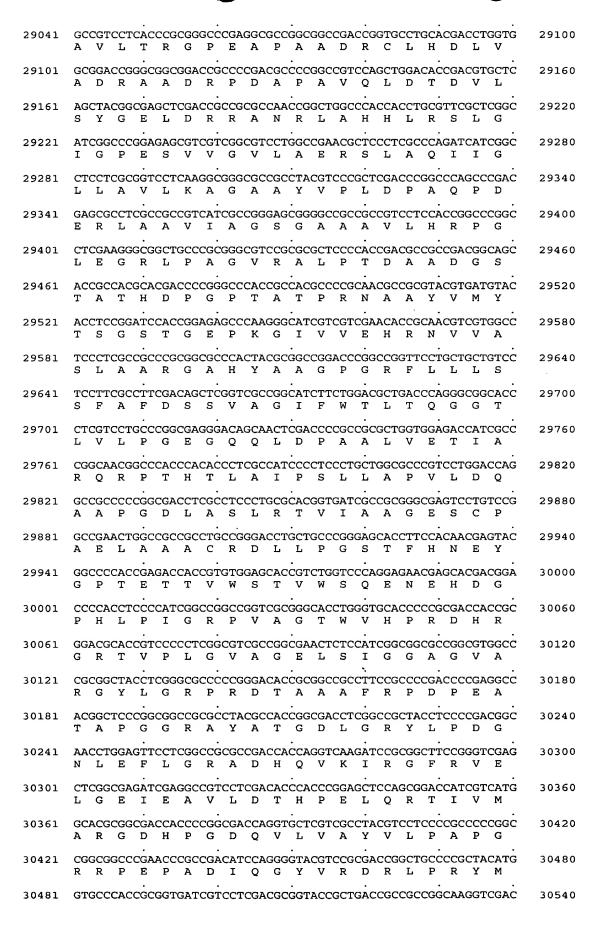


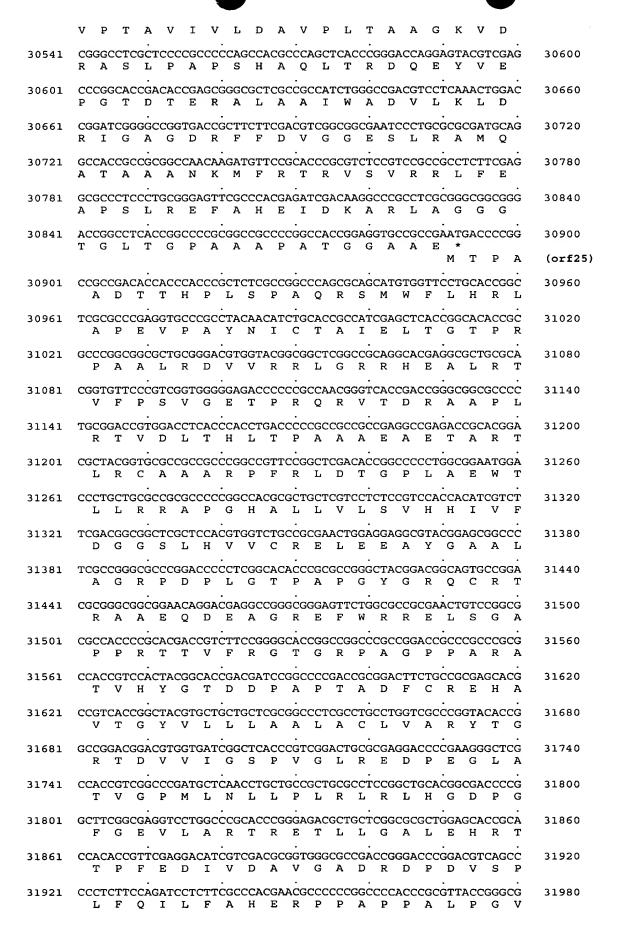


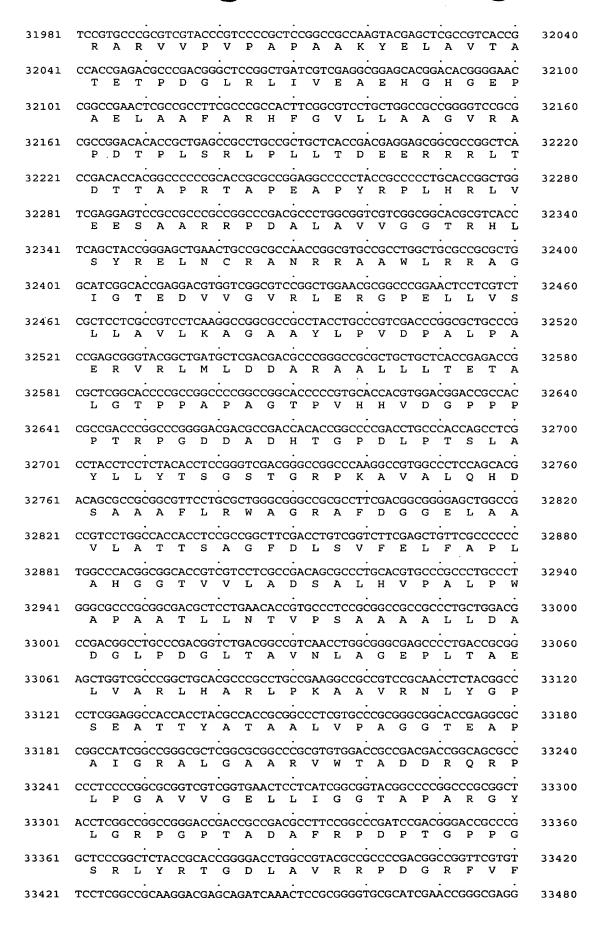


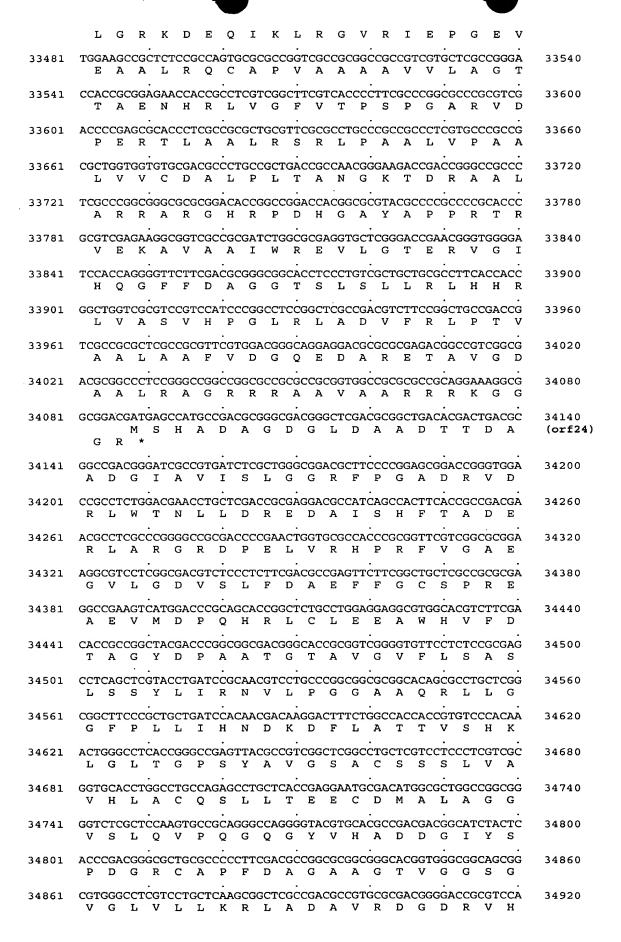


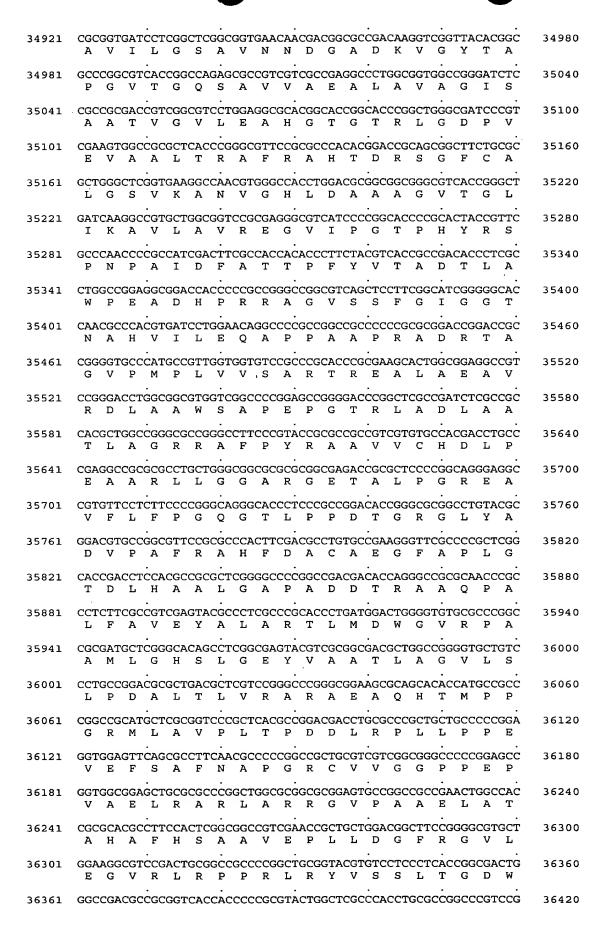


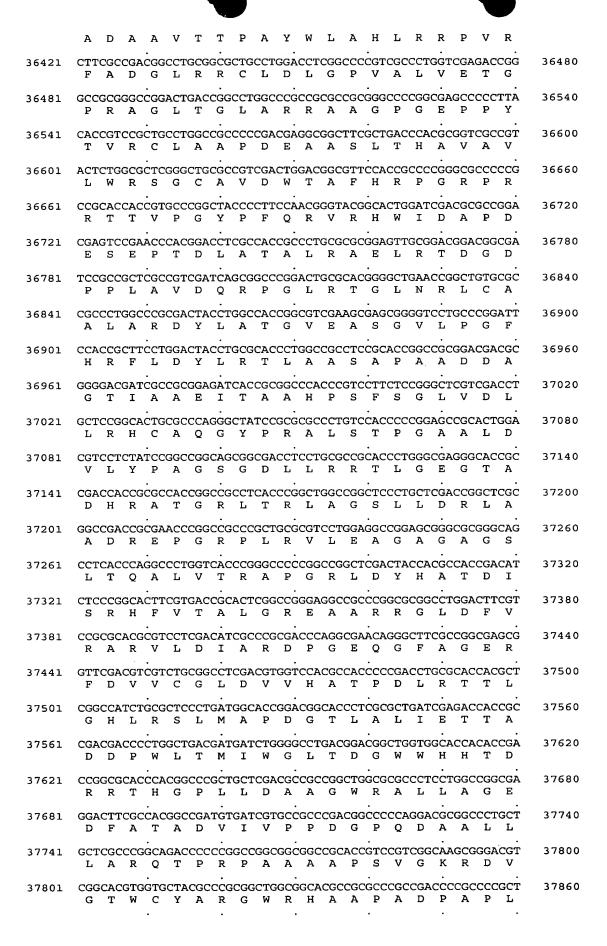


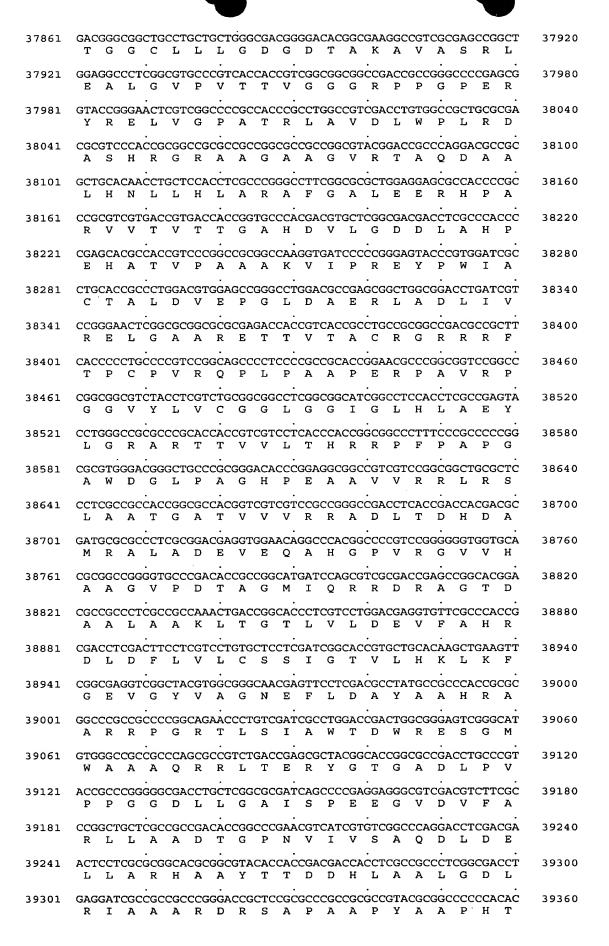


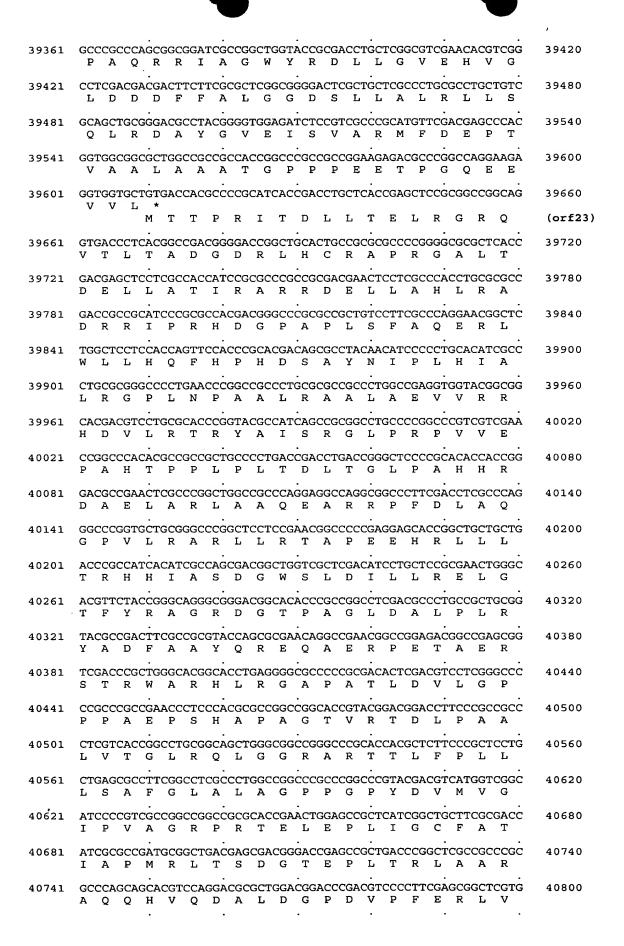


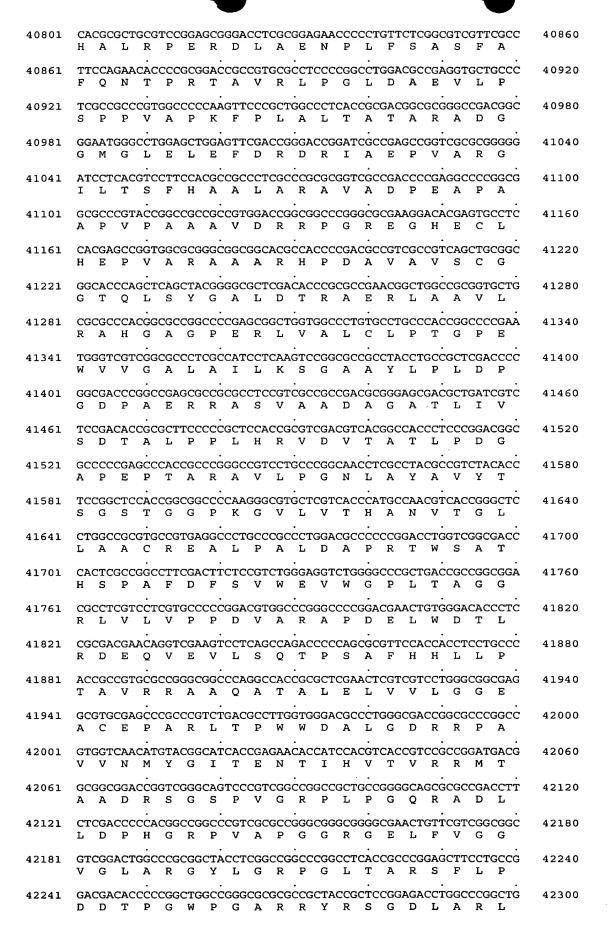




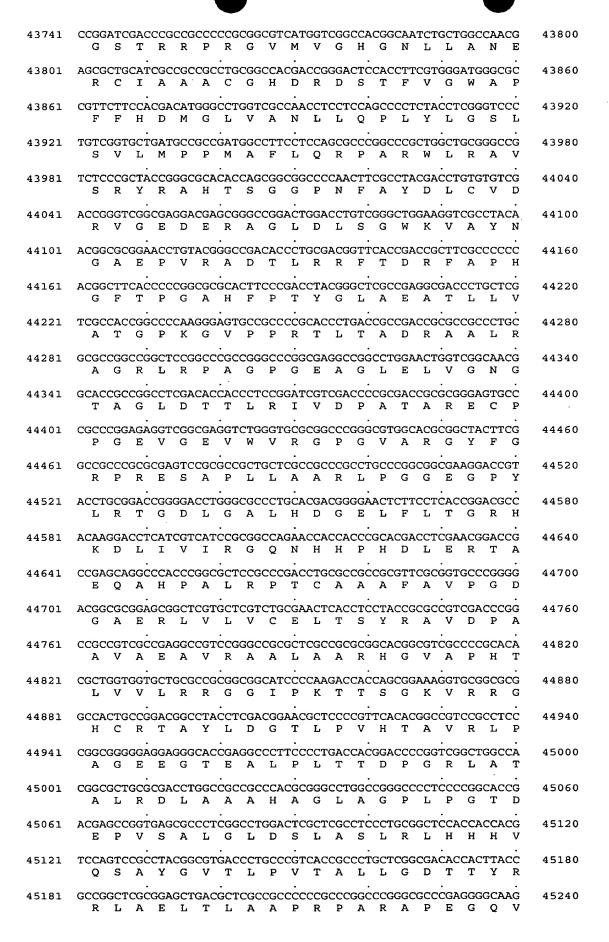


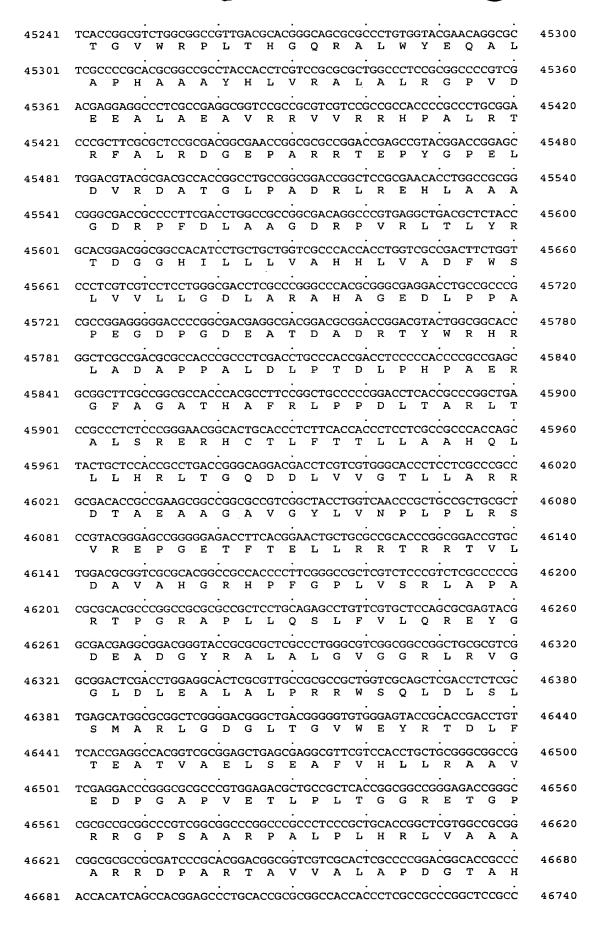


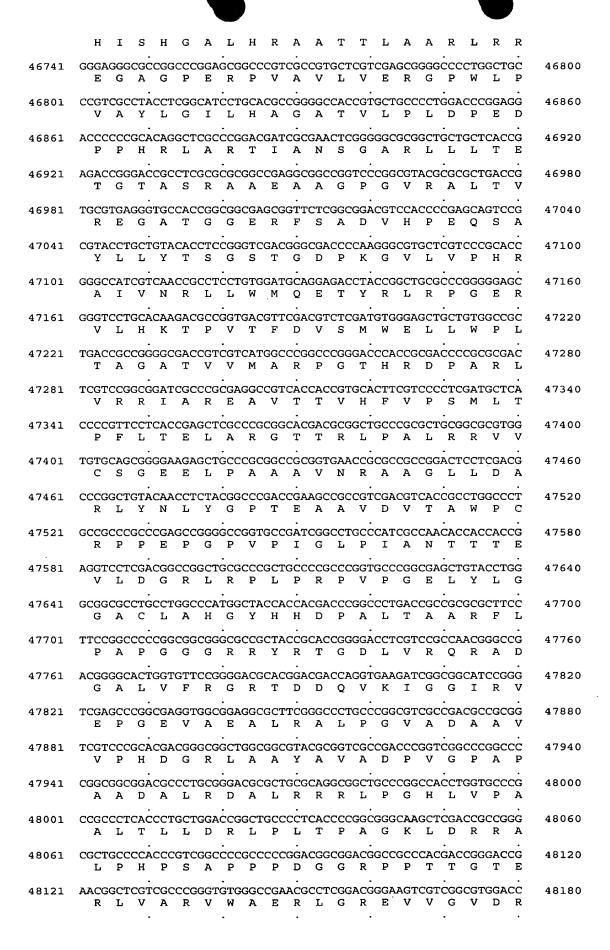


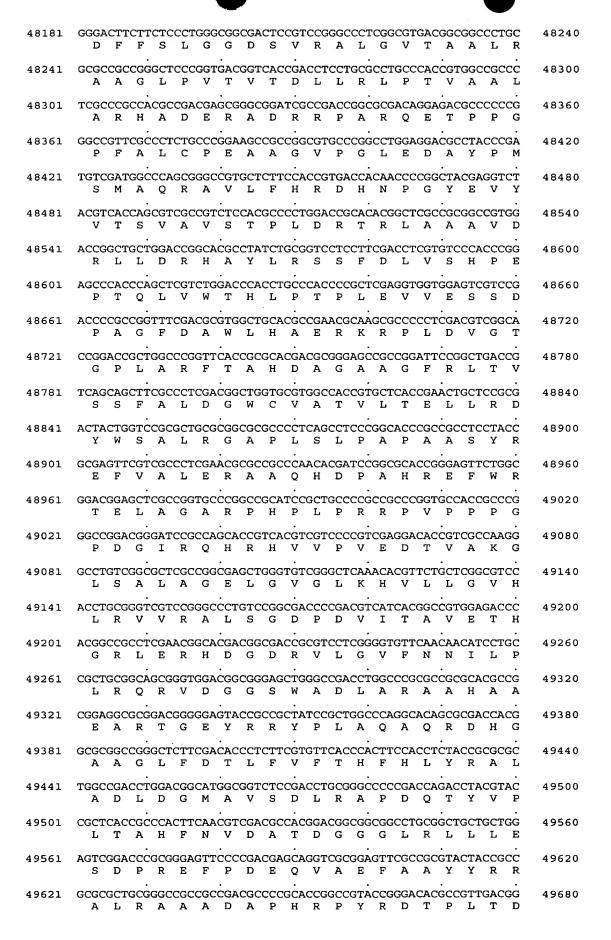


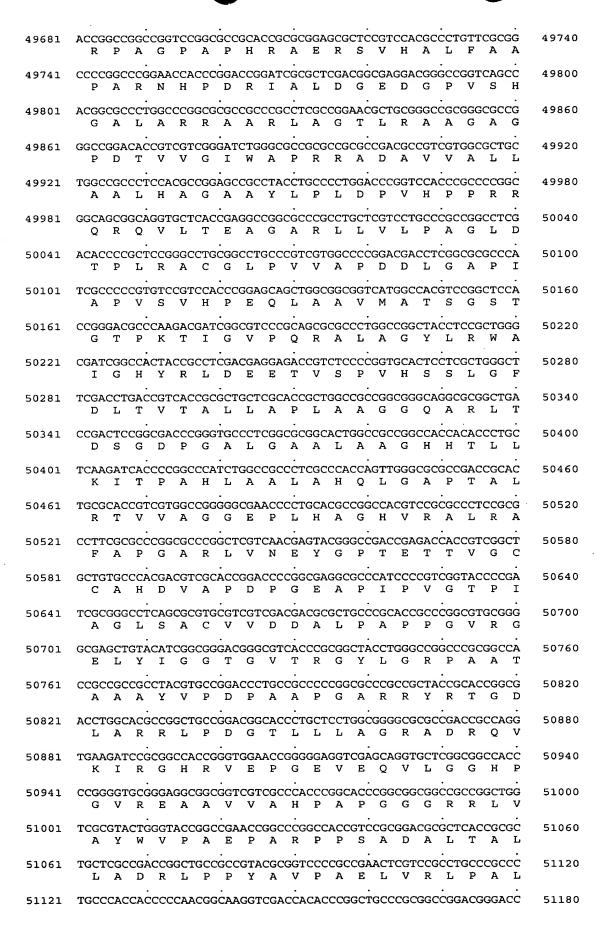
42301		42360
42361	TACCGCGTCGAGCCCGCGAGACCGAAGCCGCCGCGCGCGC	42420
42421	TGCGTGGTCGTCCACGCGGCGACGGCGACCGGCGCCATCTCGCGGCGTACGTCGTCGCCCCVVVPRGCDDGCDDRRHLAAYVVA	42480
42481	GACACCCGCGCTGCGACGGCCCGGGCTCCGCACCCACCTGGCCGAGCGGCTGCCCCGC D T R A C D G P G L R T H L A E R L P R	42540
42541	CACCTGGTGCCGGCCTCGGTGGTCTTCCTGAAGCGGATCCCGCTGACCCGCAACGGCAAG H L V P A S .V V F L K R I P L T R N G K	42600
42601	CTCGACGTGGCGGCCTTGCCCGACCCGGCCCCCCCCCCGCCCCCCCGCCCCCCCC	42660
42661	CGCACCGCGACCGAACGGACCCTCACCCGGCTGCTCGCCGCCCTCCTGAAGGCGCCACCG R T A T E R T L T R L L A A L L K A P P	42720
42721	GAGACCATCGGGACGACAACCTCTTCGACCTGGGCGGCGACTCCCTGACGGTCACC E T I G T H D N L F D L G G D S L T V T	42780
42781	CAGTTCCACTCCCGGGTGGAGGAGGTTCGCCGTGGACCTCCCGGTGCGCCGGGTCTAC Q F H S R V V E E F A V D L P V R R V Y	42840
42841	CAGGCCCTCGACATCGCGACGCTCGCCGTGACCGTGGACGACTTCCGGCGCCGCGCAAQALQALALALALALALALALALALALALALALALALA	42900
42901	CGCACCGCGGTACTGCGCCCCTCGCGGCGGCGGAGGCGATGGAACCCGGCGGTACGGCG R T A V L R A L A A A E A M E P G G T A	42960
42961	GGGGAGTCCGGCGGTAATCCGGAGGAGTCCGCCGCTACGGCGCGGGGGCCCGCCGTCGCG G E S G G N P E E S A A T A R G P A V A	43020
43021	GCGAACGAACCCGGCGCTGCGGCGCGTGAGCCCGCCGTC A N E P G A A R E S G A A P V E P A V	43080
43081	GCAGTACAGGAGTCCGCCGCTACGAAGGGGGAGCCCGCACCGCAGCGAATGAACTCGGC A V Q E S A A T K G E P G T A A N E L G	43140
43141	GCTGAGGCACGGGAGCCCGGCACCGCACCCCACCC A E A R E P G T A A Q E P G T D P R P P	43200
43201	GCCGCCACACCGCAGGACCCCCGCACCACCACCGCAGGAAGGACAGCCGTGCCCGCGTCCC  A A T P Q D P R T T P Q E G Q P C P R P	43260
43261	GAATGAGCCGGCCGGCATCGTCGACATCGCGCGCCGTCACGCCGAGCGCACCCCCG  M S R P A G I V D I A R R H A E R T P A E *	43320 (orf22)
43321	CCCGTCCCGCGTACGCGTTCCTGCCCGACGGCGAGACGGAGAGCGTCCGCTTCTCCTTCG R P A Y A F L P D G E T E S V R F S F A	43380
43381	CCGACATCGACCGGGGGCCGCGCCGTGGCCGCCGTCCTCCAGGACCGCGGCCTGGCCG D I D R R A R A V A A V L Q D R G L A G	43440
43441	GGGAGCGGGTCCTGGTCGCCTATCCCTCCGGGCCCGAGTACGTCCAGGCGTTCCTGGGCT E R V L V A Y P S G P E Y V Q A F L G C	43500
43501	GCCTGTACGCGGGCGTGGTCGCCGTCCCCTGCGACGAGCCGCGCCCGAGCGCGG L Y A G V V A V P C D E P R S G P S A E	43560
43561	AACGGCTCGCCGGGATCCGCCGACGCCCGCCCCGCCCTGACCGCCGCGCCCCCCCC	43620
43621	CCGAGGCCGGGCTCGCCGGCCTGGCCACCCTGGACGTGGCCGGCGTCCCCGACTCCGCCG E A G L A G L A T L D V A G V P D S A A	43680
43681	CCGGGGCCTGGACCGACCCCGTCGCGGACCGCCTGGCCTTCCTCCAGTACACCT G A W T D P V A G P D A L A F L Q Y T S	43740

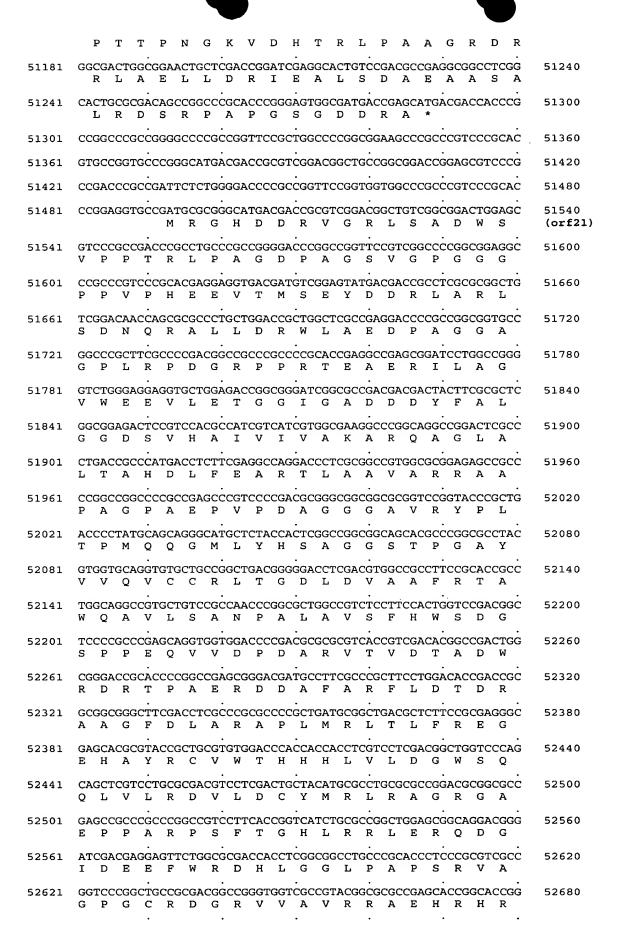


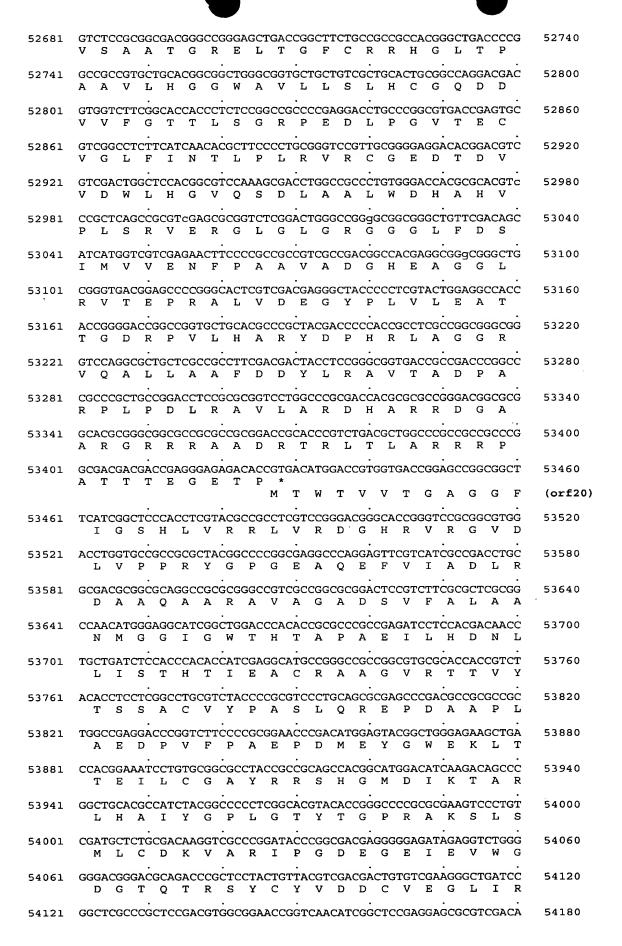


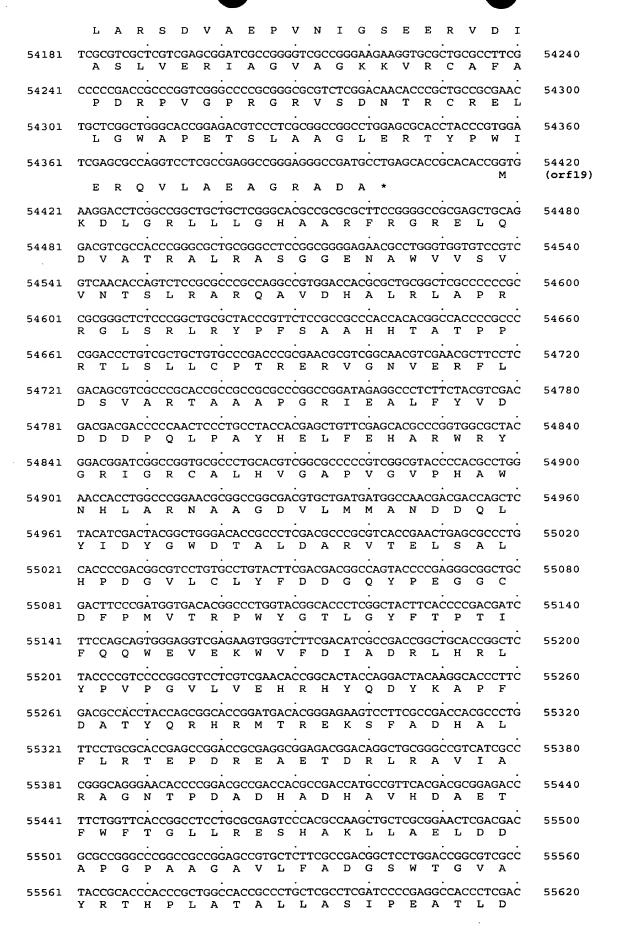


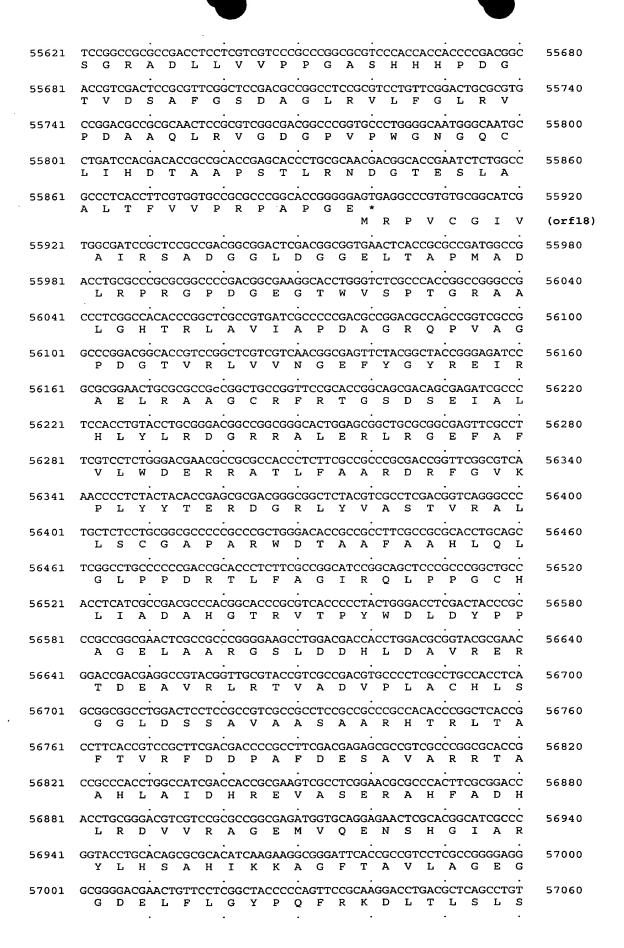


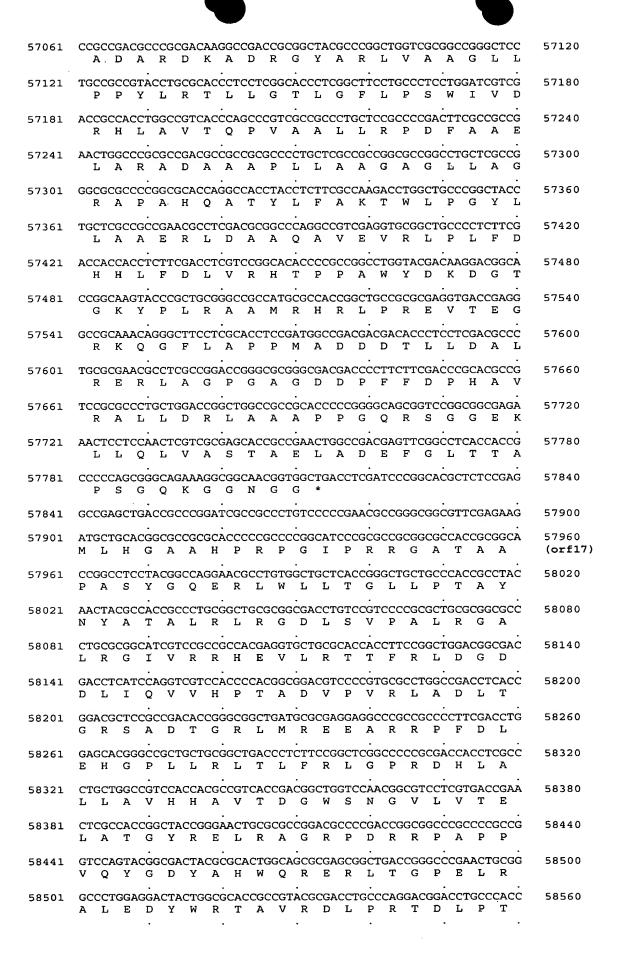


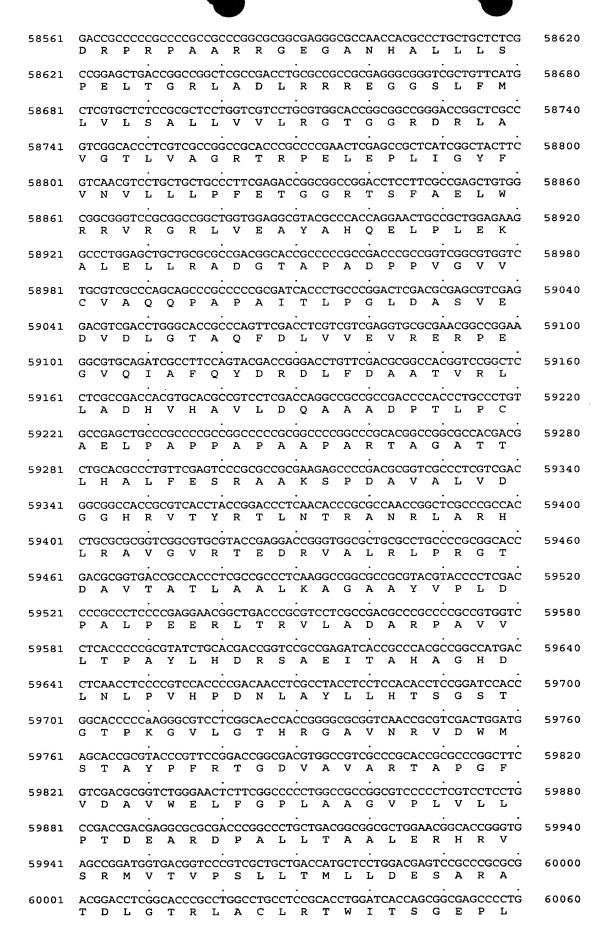


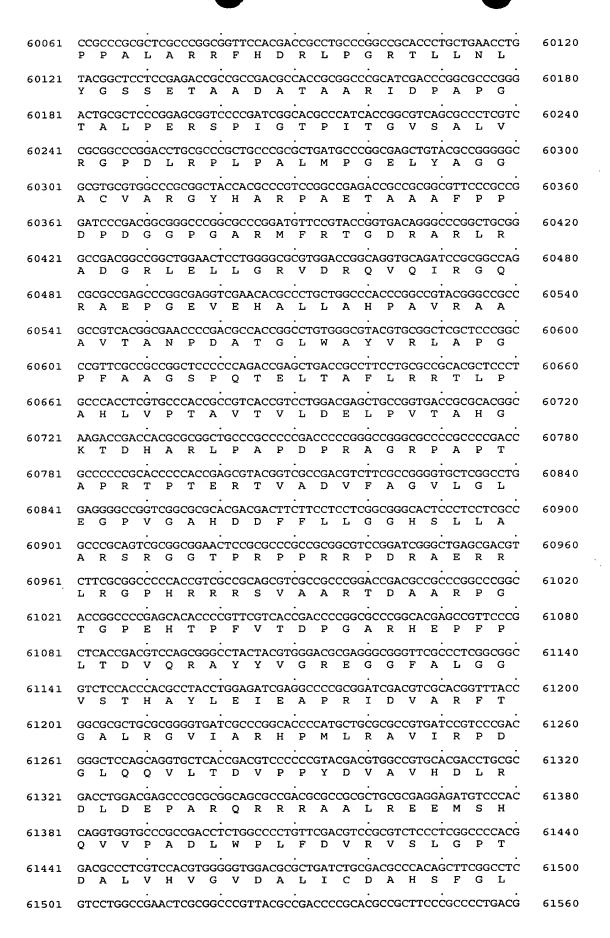


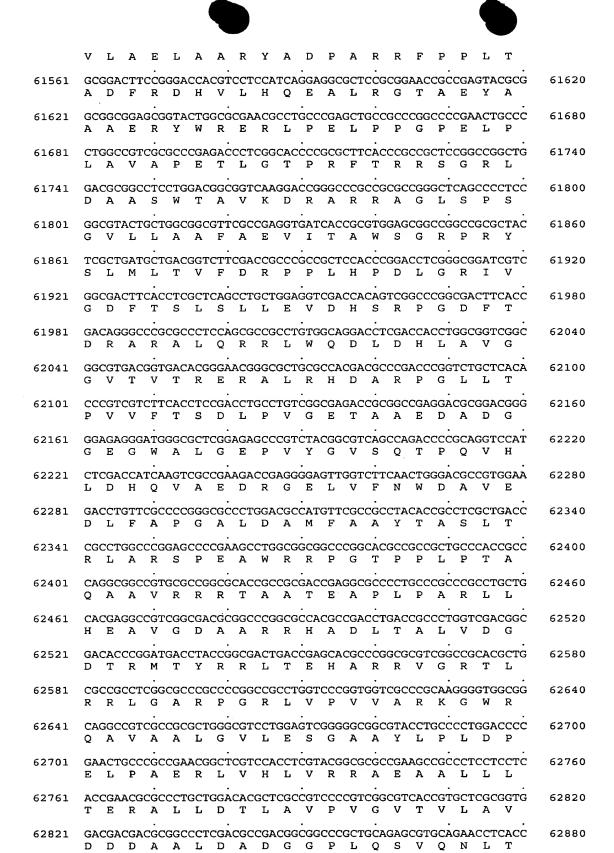














62940

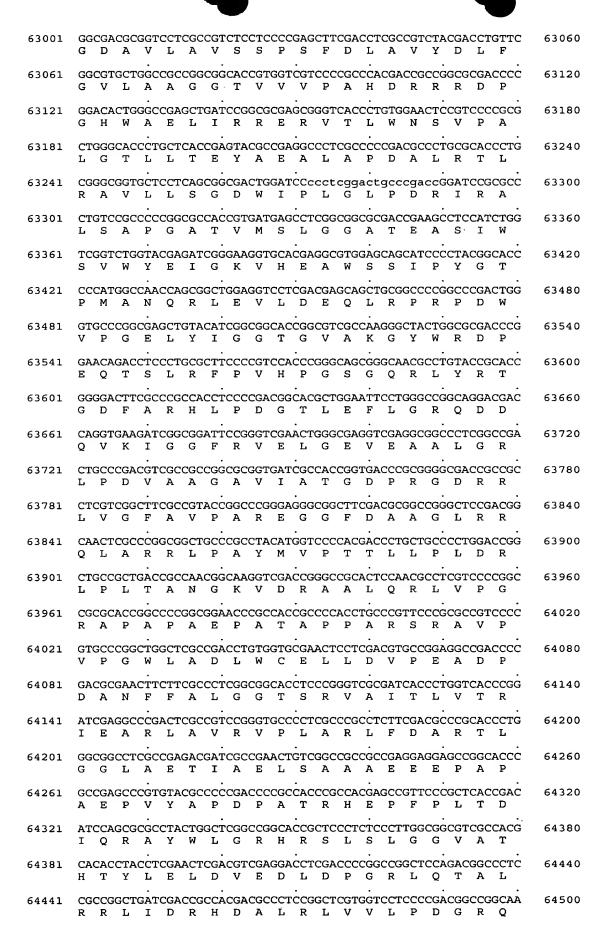
63000

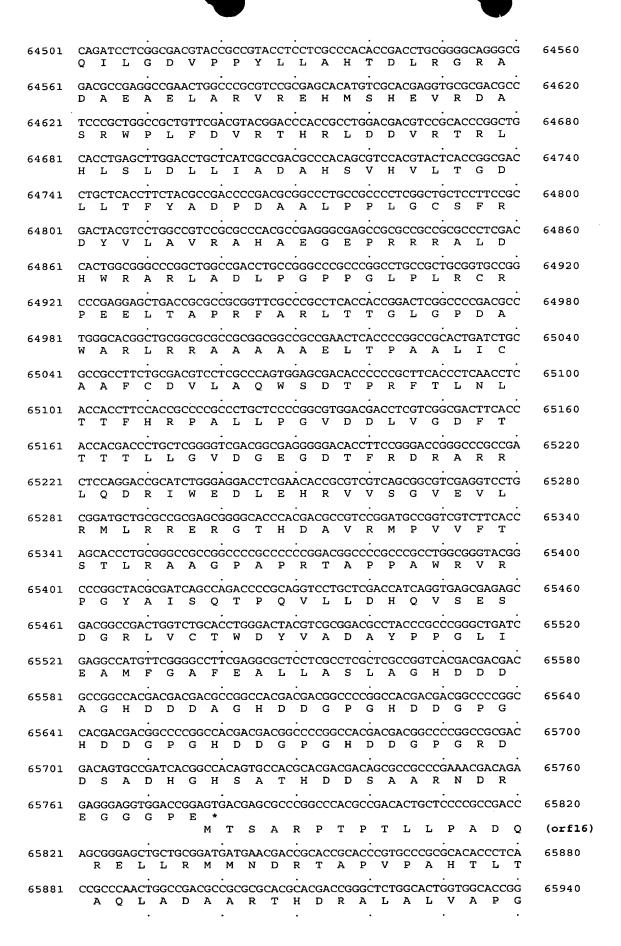
62881 GACCTGGCGTACACCATCTTCACCTCGGGCTCCACCGGCGAACCCAAGGGCGTCATGATC

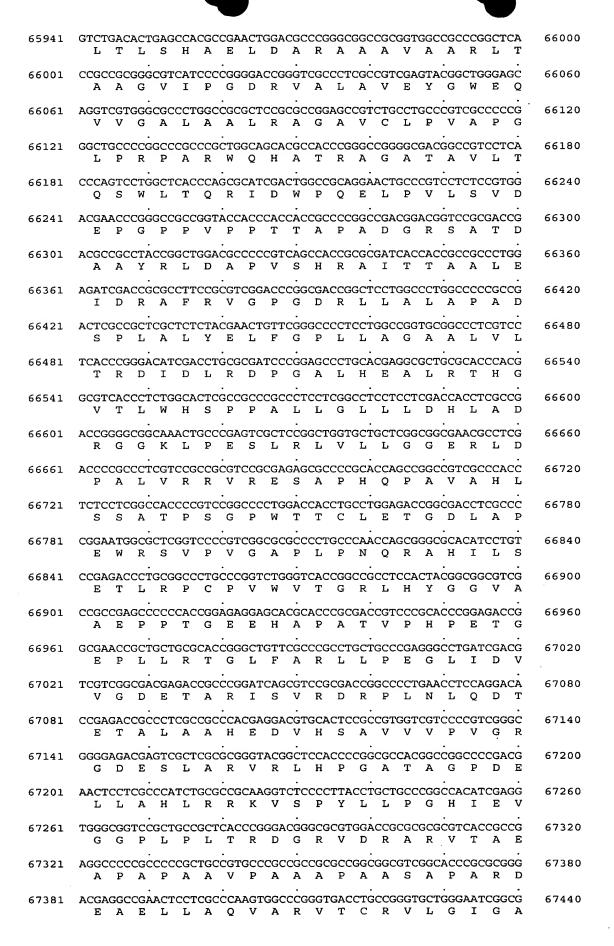
62941 GACCACCTCGGCGCGCCAACACCCTGGAATGCGTCAACCGCCGCTTCGGCACCGGCCCC

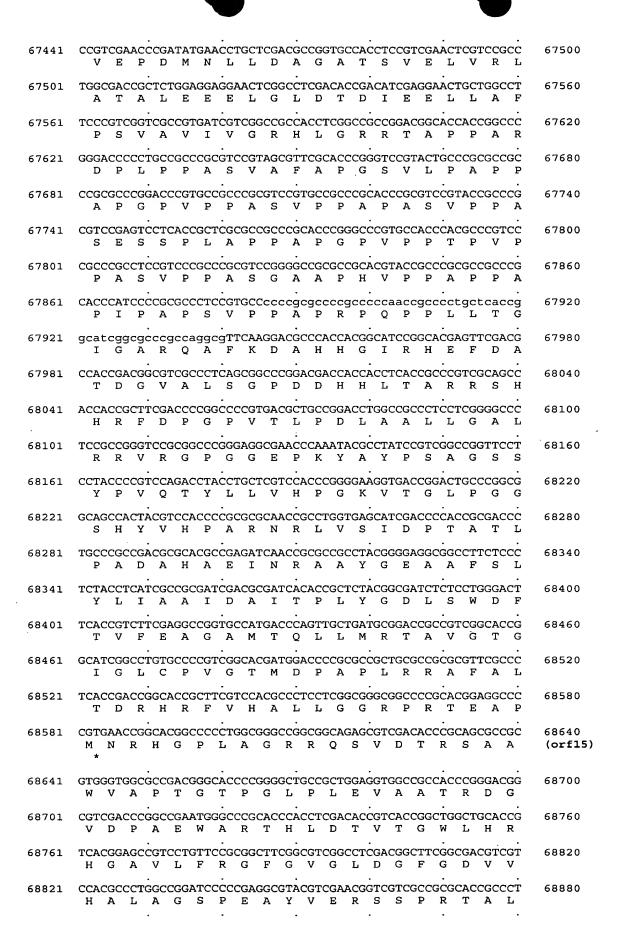
DLAYTIFTSGSTGEPKGVMI

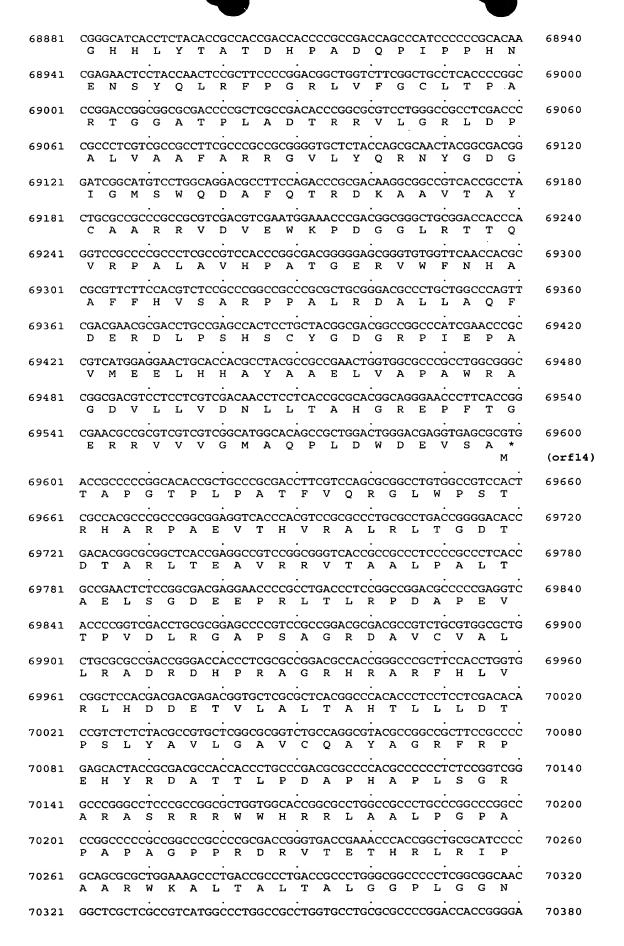
D H L G A A N T L E C V N R R F G T G P

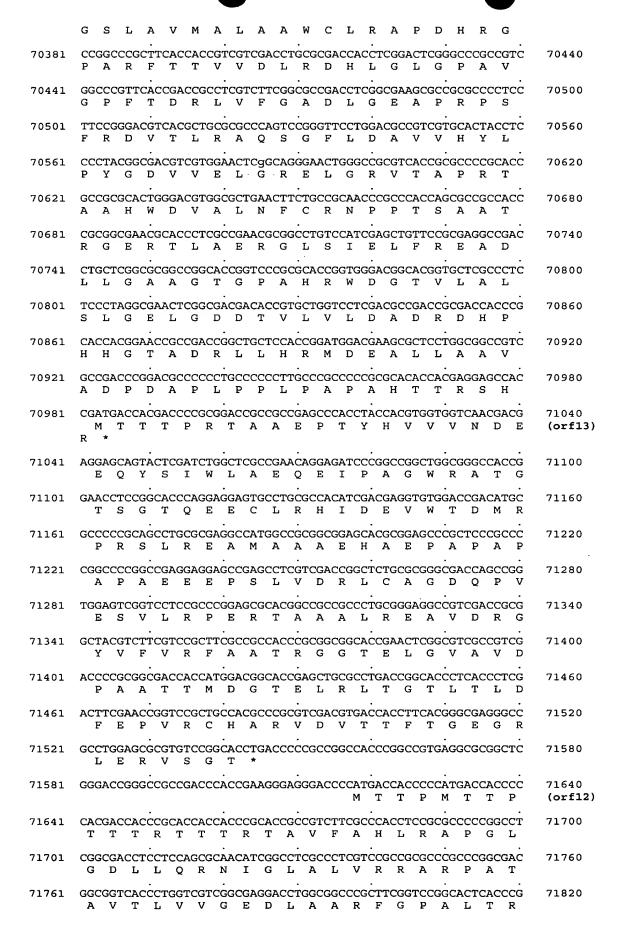


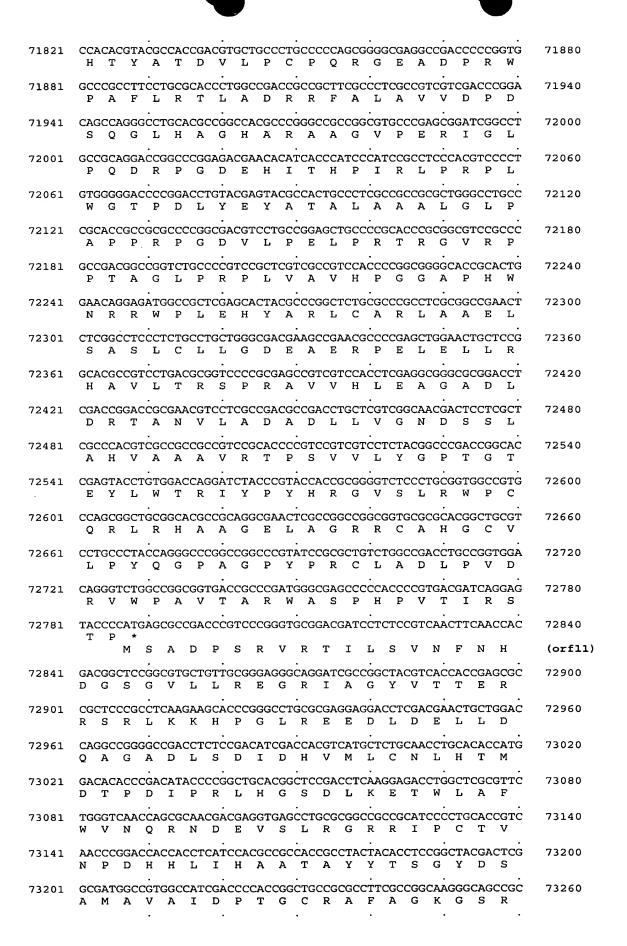


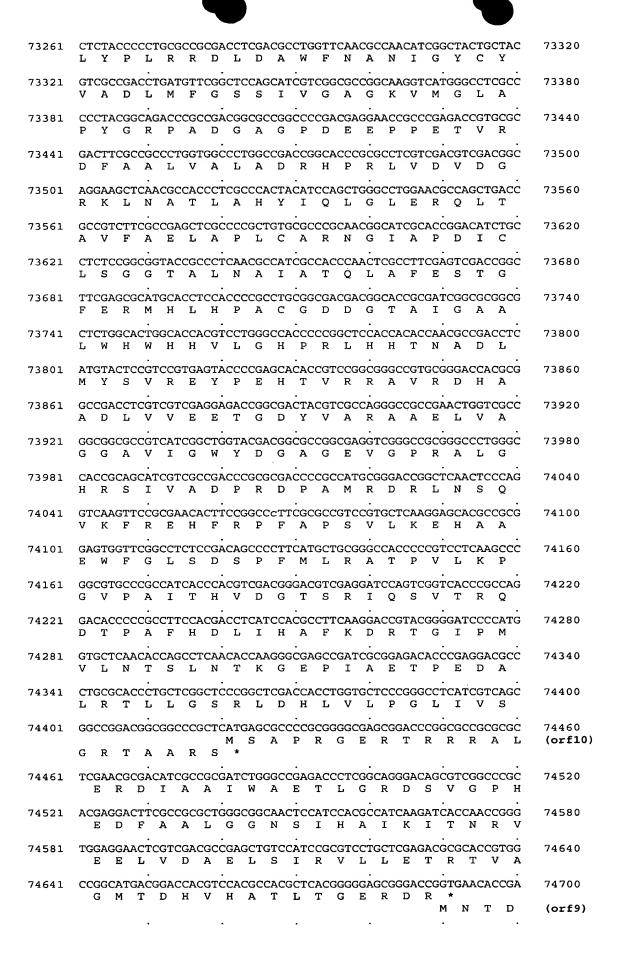


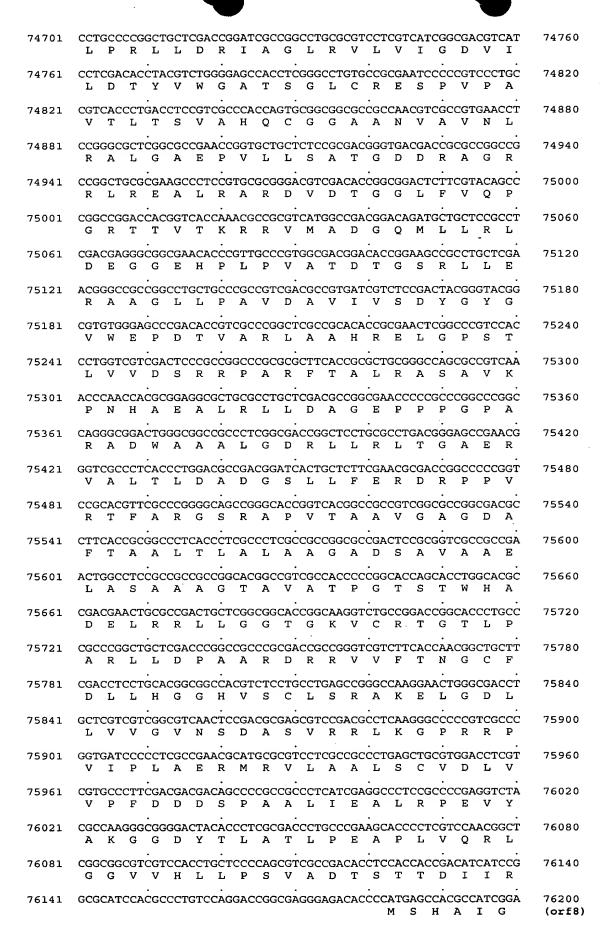








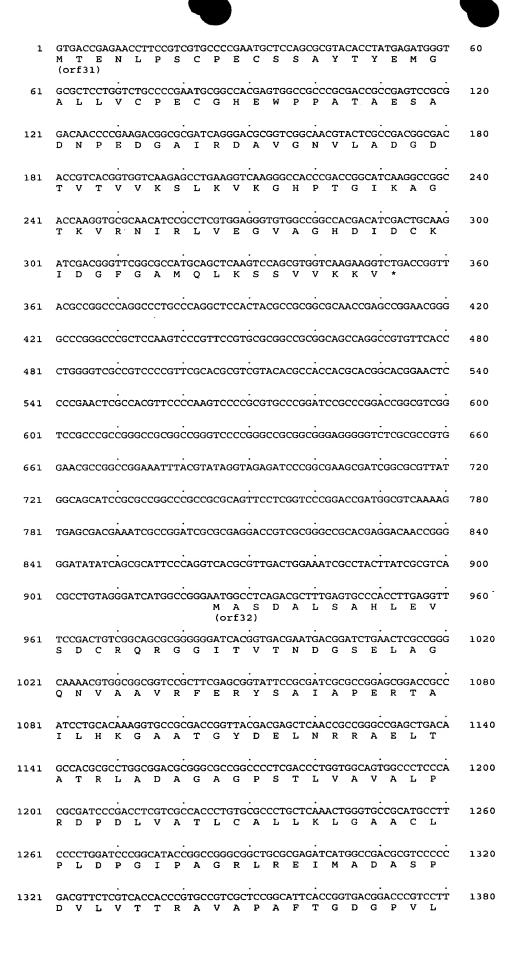


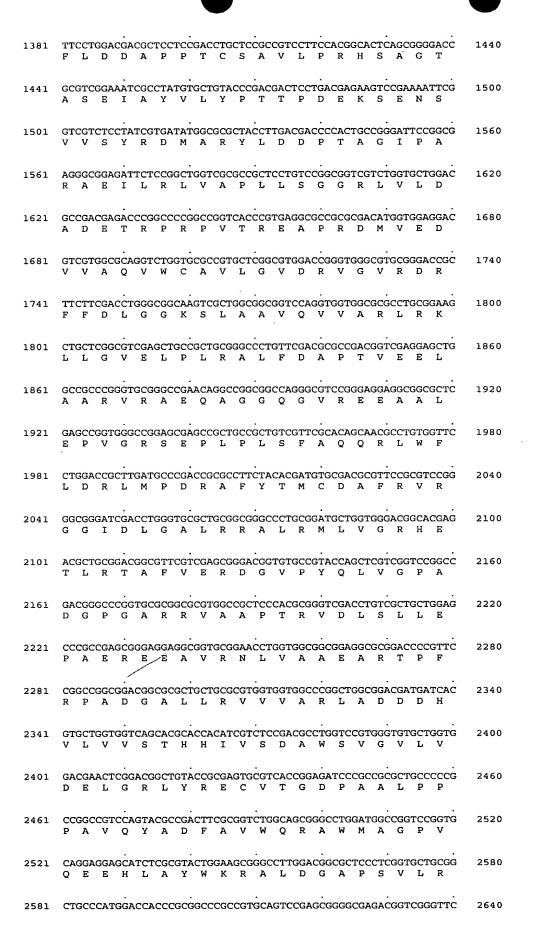


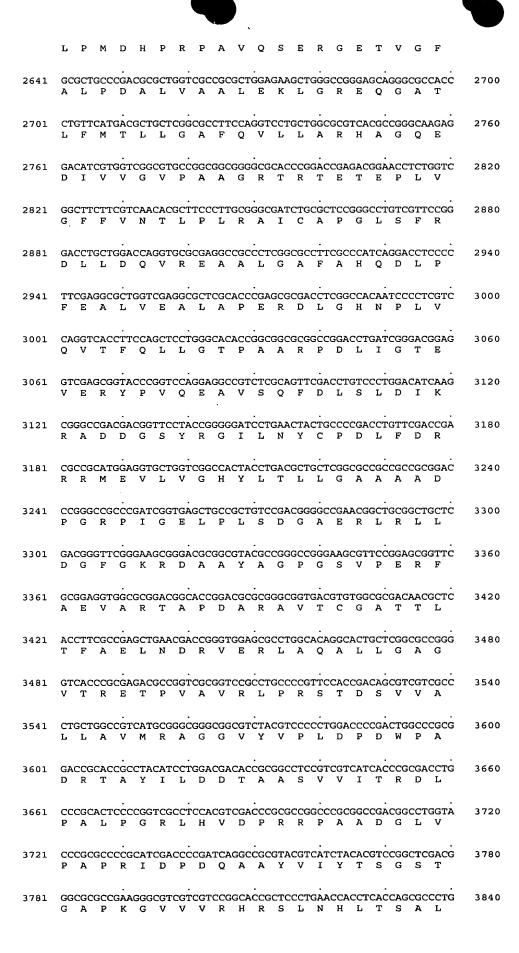
## RIHALSRTGEGDTP 76201 CCGAGCCGGCTGATCCCCGCCATCCGCGAAGCGCTCGGGGACGAAGGACCCCCGGCTC 76260 P S R L I P A I R E A L G D E K D P R L 76261 GCCCTCTACGTCCACGTCCCCTTCTGCTCCTCCAAGTGCCACTTCTGCGACTGGGTCACC 76320 ALYVHVPFCSSKCHFCDWVT 76321 GACATCCCGTCGCACGCCTGCGCGGGCACGCCGGGAACGCTCGCCCTACGTCACCGCC 76380 DIPVARLRGDSRERSPYVTA 76381 CTCTGCGACCAGATCCGCTTCTACGGCCCCCAGCTCACCCGGCTCGGCTACCGCCCCGAG 76440 L C D Q I R F Y G P Q L T R L G Y R P E 76441 GTCATGTACTGGGGCGGCGCCCCCCCCCCGGCTCACCGGCGACGAGATGACGGCCGTC 76500 V M Y W G G G T P T R L T G D E M T A V 76501 CACCAGGCCCTCGACGCCTTCGACCTGACGGGACTCCGCCAGTGGTCGGTGGAGAGC H Q A L D D A F D L T G L R Q W S V E S 76620 T P N D L D P A T L D T L R G L G V T R 76680 V S V G V Q S L N P Y Q L R K A G R A H 76681 TCGCGCGAACAGGCCCTGGCCGCCGTCCCCCTGTTGCGCCGCGCCGCGCATCGACGAGTTC 76740 S R E Q A L A A V P L L R R A G I D E F 76741 AACGTCGACCTGATCGCCGGCTTCCCCGGCGAAGCCGTCGAGTCCTTCGAGGAGACCCTG 76800 N V D L I A G F P G E A V E S F E E T L 76801 CGCACCGTCTCGCGCTCGACCCGCCGCACGTCTCCGTCTACCCCTACCGCGCCACCCCC 76860 R T V L A L D P P H V S V Y P Y R A T P 76861 AAGACGGTCATGGCCATGCAGCTCGACCGCGAGTTCGTCGAGGCCCGGAACCGGGACGGC 76920 K T V M A M Q L D R E F V E A R N R D G 76921 ATGATCGACGCCTATGAACGGGCCATGGCCGCCTCGGCGCCGCCGGCTATCACGAGTAC 76980 MIDAYERAMAALGAAGYHEY 76981 TGCCACGGCTACTGGGTGCGCGACGCGCCACGAGGACCAGGACGGCAACTACAAGTAC 77040 C H G Y W V R D A R H E D Q D G N Y K Y 77041 GACCTGGCCGGCGACAAGATCGGCTTTGGCAGCGCGCCGAATCGATCATCGGTCACCAC 77100 D L A G D K I G F G S G A E S I I G H H 77101 CTGCTCTGGAACGAGAACAGCGCCTACGCCCGCTACCTGCTCGCCCCCCGCGAGTTCTCC 77160 LLWNENSAYARYLLAPREFS 77161 GCCGCCCACCGGTTCACCACCGCCGAACCCGACCGCCTGACCGCCCCCGTCGGCGCCGCC 77220 A A H R F T T A E P D R L T A P V G G A 77221 CTGATGACCCGTGAAGGCGTGGTCTTCGCCCGCTTCCGCAGACTGACCGGCCTGGACTTC 77280 LMTREGVVFARFRRLTGLDF 77281 GCGGACGTCCGCGCACACCGTACTTCCGCCAGTGGTTCGAGCTCCTGGAGCGCTGCGGC 77340 A D V R A T P Y F R Q W F E L L E R C G 77341 GGCCGCTTCGTCGAGACGCCGTACAGCCTCCGCCTGGAGCCGTCCACCATCCACCGCGCC 77400 G R F V E T P Y S L R L E P S T I H R A 77401 TACATCACCCACCTCGCCTACACCATGGCCCATGGCCCCGGAACGCGCCTGA 77457 YITHLAYTMAHGLAPERA\*

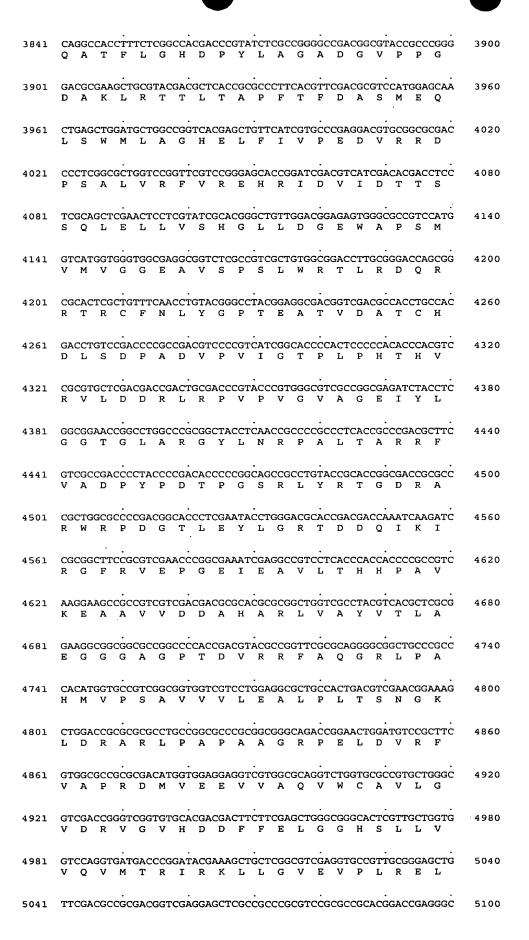
## SEQ ID NO: 2 ORFS BLM gene cluster ORFs 31-40

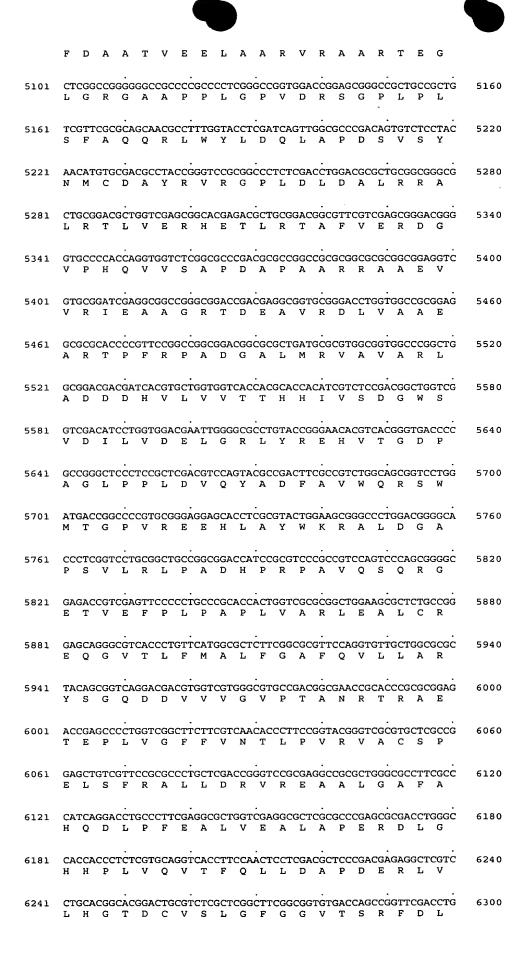
(notice this part is on the reverse strand and the last nucleotide (18660) is the first (1) on the whole cluster of 77457 bp. Also the last orf (40) is incomplete and contains frame shifts)

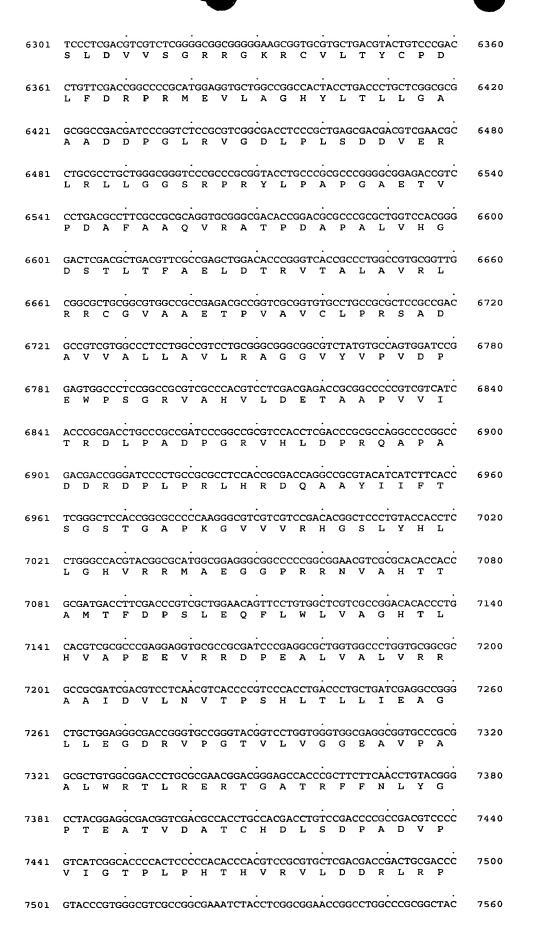


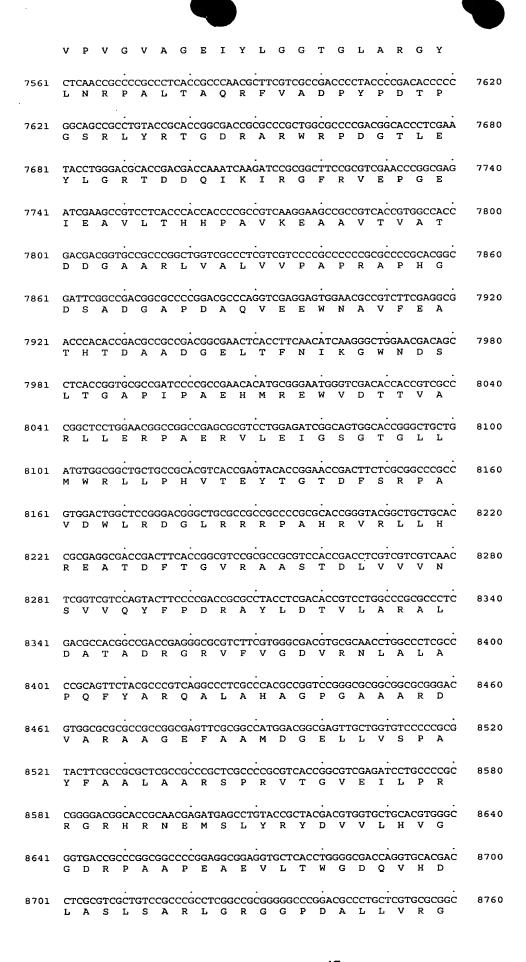


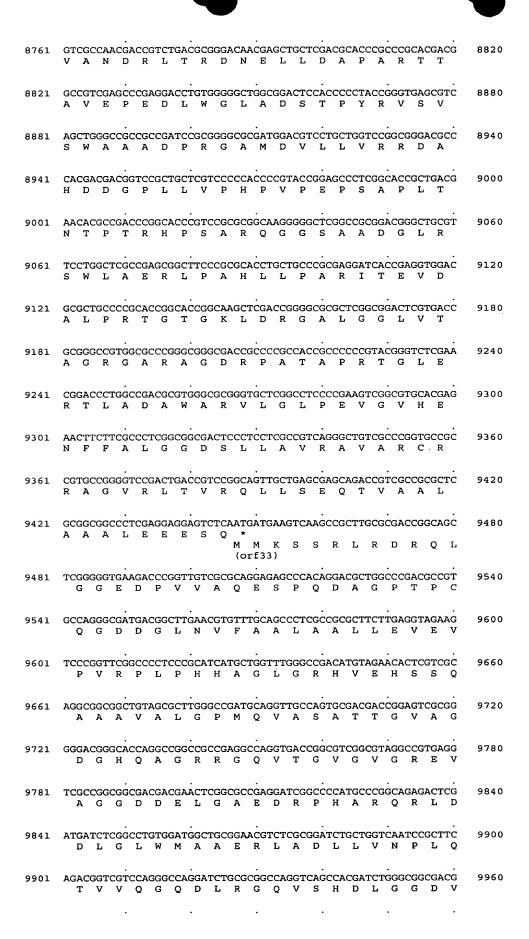


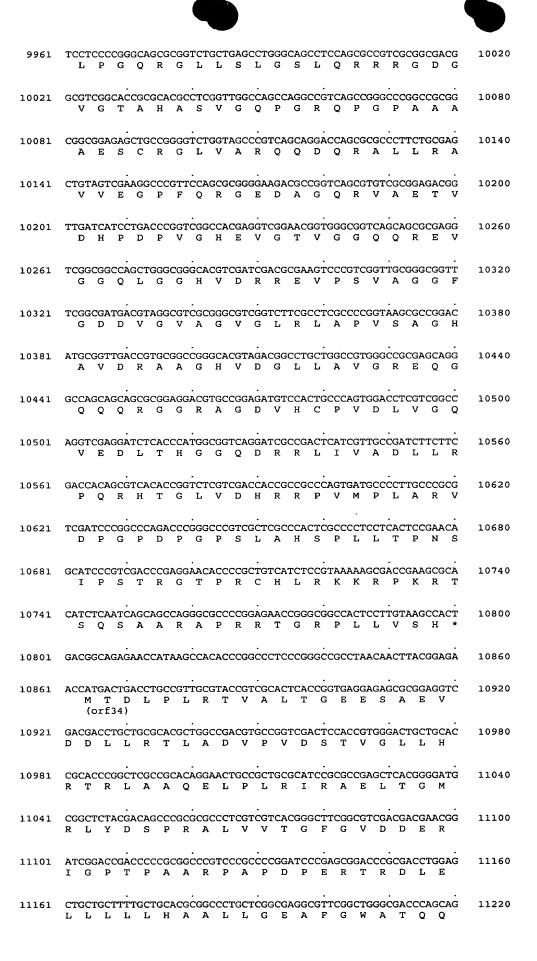


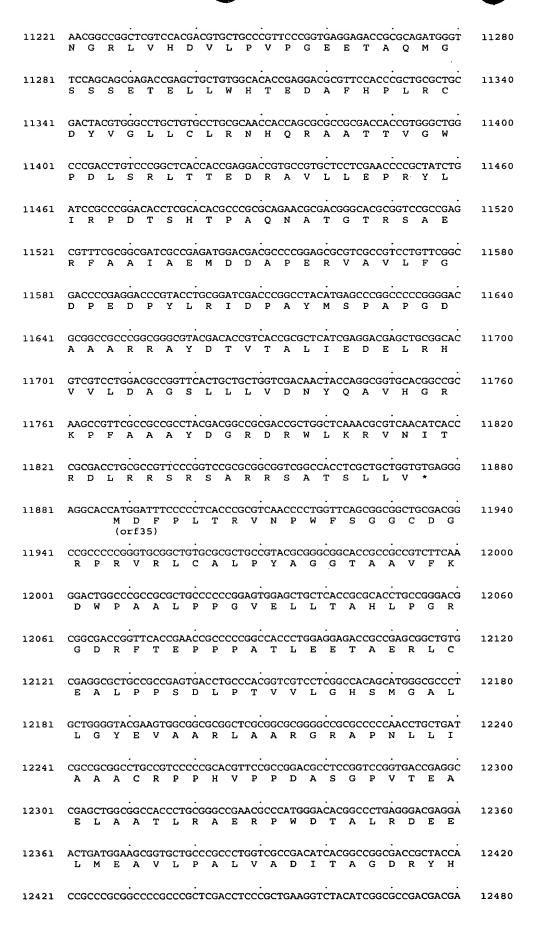


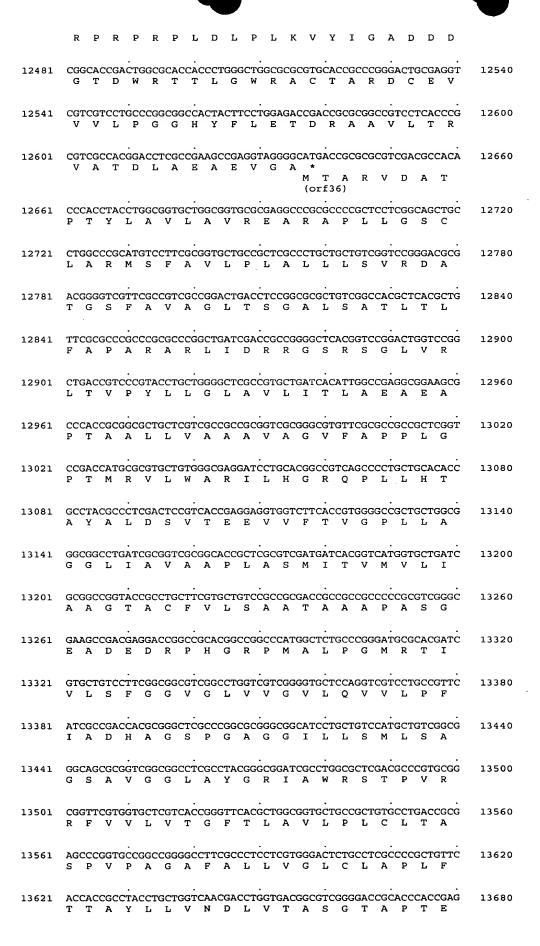


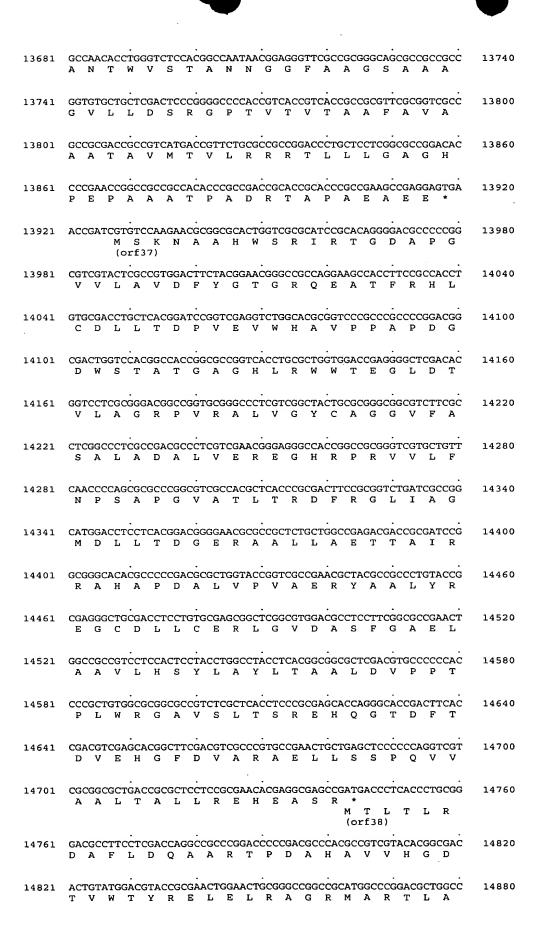


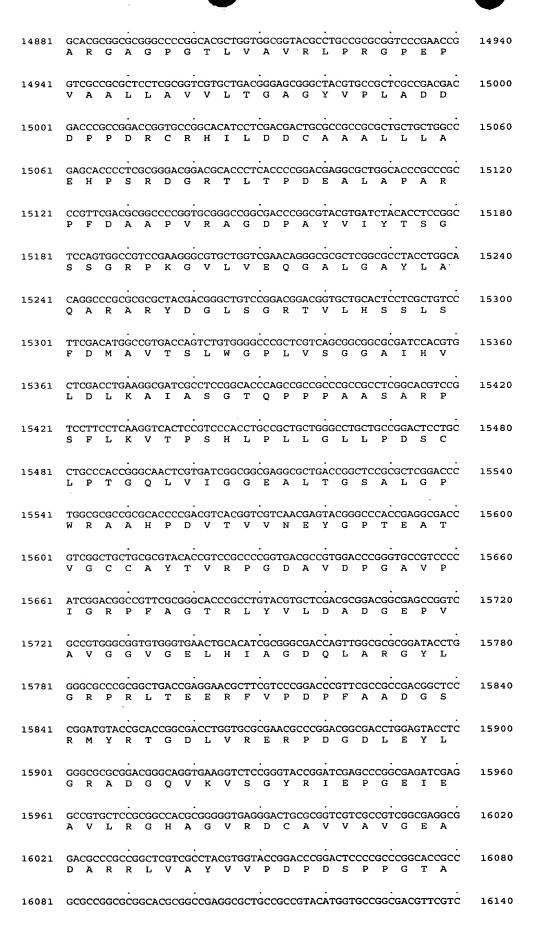


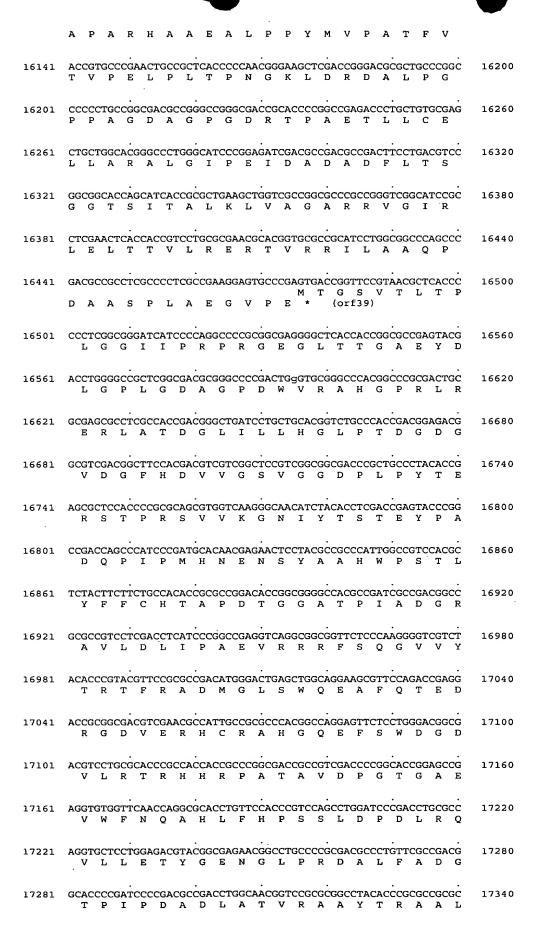


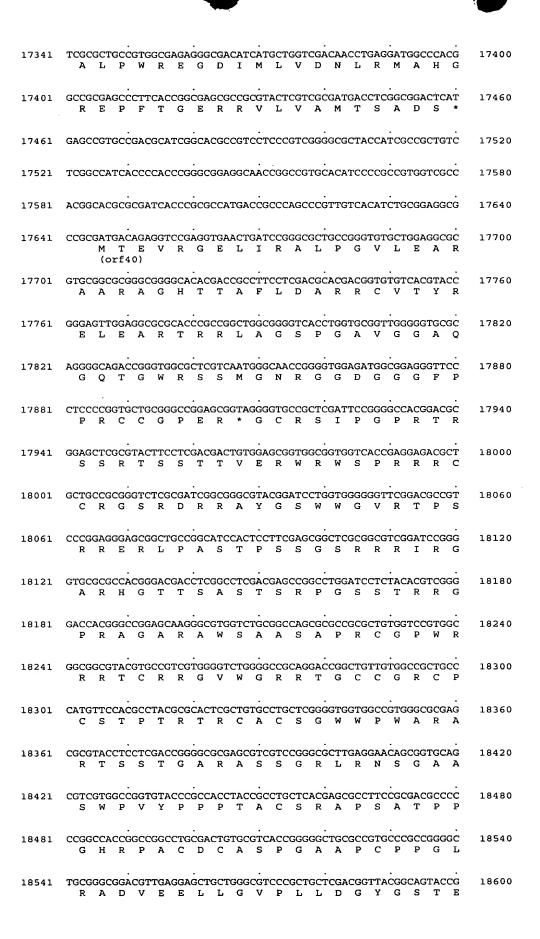
















18601 AGACCTGCGGCAAGATCACGGTTGAGCGGCTCGGCGGCGCTCCCGGGAGGGCGGTTGCCGGG 18660
T C G K I T V E R L G G S R E G G C R

## SEQ ID NO: 3 BLM gene PPTase ORFS 41

1	GGATCCTGCGCTACCCGGACTTCGCCCAGTGGTGCGGCACCGAGCTCACCGCCGACTGGCACGTCCGCCTTCCGGGCCGCC	80
81	GCCGCGGTCTACGGGCATCTGCACATCCCCCGCGTGACCCGGGTACGACGGCGTCCGCTTCGAGGAGGTGTCGGTCG	160
161	CCCGCGCGAGTGGCGGCCGCGCGGCGGCGGGGGGGCCGCTCGGCAGATCCTGCCCCAGCCCGTCGACGAGCCGGGAGCCC	240
241 1	TCTGGTGATCGCCGCCCTCCTGCCCTCCTGGGCCGTCACCGAACACGCCTTCACCGACGCCCCGGACGACCCGGTGAGCC M I A A L L P S W A V T E H A F T D A P D D P V S L	320 26
321 27	TCCTCTTCCCCGAGGAGCCGCCCACGTCGCCCGCGCCGCCCCCAAGCGCCTGCACGAGTTCGCCACCGTCCGGGTGTGC  L F P E E A A H V A R A V P K R L H E F A T V R V C	400 52
401 53	GCCCGCGCCCCCCGGCCGGCCGGCCGGCCGGCCGAGCGGCG	480 79
481 80	CGGGGTGGTGGGGAGCATGACGCACTGTC $\lambda$ GGGCTTCCGGGGCCGCGGCGCGCGCGCGCGCGCGCGCGC	560 106
561 107	GGATAGACGCCGAGCCGACGGGCCCTCCCGGACGCGCTCCTCGCCATGGTCTCGCTGCCGTCCGAGCGCGAGTGGCTC I D A E P N G P L P D G V L A M V S L P S E R E W L	640 132
641 133	GCCGGACTGCCGCCGGCCGGACGTGCACTGGGACCGGCTGCTGTTCAGCGCCAAGGAGAGCGTCTTCAAGGCGTG A G L A A R R P D V H W D R L L F S A K E S V F K A W	720 159
721 160	GTACCCGCTGACCGGCCTGGAGCTGGACTTCGACGAGGCCGAGCCGGGACGCCGGGACGTTCACGGCCC Y P L T G L E L D F D E A E L A V D P D A G T F T A R	800 186
801 187	GGCTGCTGGTGCCGGGACCGGTGGTCGGCCGGCCGGCGGCGGCGGGCG	880 212
881 213	GTCGTCACGGCCATCGCCGCGGCGGCCGGCCGGCCGGCGGGAGGAATCGGCGGAAGGGAAGGAA	960 239
961 240	GGACGACCGGACCGCCGTAAACCGCCCCGAACACCGGCGTGGCGCCGACCGTGTCGGGGGGCGCCACGAACG D D R T A V P *	1040 247
1041	GGCGCCGGCCGGCGGGCCCTCCGCCGTGCGGAGCCGGAGGCCCGGACGCCCGGTGTCGTCGGATACGTGCGTC	1120
1121	AGTCGGCGACGCAGACGTTGCCGTTGGTCGAGTTGACCAGCCCGACGATGTCGATGGTGTTGCCGCAGAGGTTGATGGGG	1200
1201	ATGTGGACGGGGATCTGGATGACGTTGCCCGAGACGACGCCCGGGGGAGCCGACGGCCCCCTTGGCGTTCGAGTCGGC	1280
1281	GAGGGCGGTGCCGGAGACGCCGGCGAGCGCCGTGCCCACGGTGGCGGTGAGGGCCGCTGCCTTGGCGATTCGTGACATGG	1360
)1361	GGTGACACCTTCGTTCGGTCTGACAGGGTCGAGCTCACGGCCTCTGACGGCCGGGAGCCCGGATCAACGCCCGATCACCC	1440
1441	CGAAGGTTTCGAATCGTGCGGCGGACGGGTGACCGGCGCCCGAACGGCCTCGCCGGGACGCCCCGGAAGGTGCCATGACGTC	1520
1521	CGTGCGCCATCTGTACAGCCCGGTCCCGCGCGCGCGTACAAGGGACGGAC	1600
1601	GGGGAGGCCATGAGCCGGATCGCGATCGTCGGGGCGGGTCAGGCCGGACTGCATCTGGCGCTGGGGGCTGCTGGGGGCGGG	1680
1681	GAGCGGCTCTTCCCGTCACGAGGTGCTGCTCGTGTCCGACGGGACGACGAGATCCGCGCCGGGCGGG	1760

/

1761 C 1761